DOCUMENT RESUME

ED 301 576 TM 012 435

AUTHOR Perkins, Kyle
TITLE A Comparison of Instructional Sensitivity Indices.
PUB DATE 84
NOTE 23p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Bayesian Statistics; Comparative Analysis; *Criterion Referenced Tests; Difficulty Level; *English (Second Language); Grammar; Higher Education; Language Proficiency; *Language Tests; *Latent Trait Theory; Test Construction; Test Items; Undergraduate Students
IDENTIFIERS Classical Test Theory; *Instructional Sensitivity; *Michigan Test of English Language Proficiency

ABSTRACT

Instructional sensitivity indices are compared from four perspectives: (1) criterion-referenced testing; (2) classical test theory; (3) item response theory; and (4) Bayesian theory. Instructional sensitivity is defined as the tendency of a test item to vary in difficulty as a function of instruction. The instructional sensitivity of the items in the grammar subtest of the Michigan Test of English Language Proficiency was assessed as an example. Subjects (N=18) were enrolled full-time in proficiency level 3 at the Center for English as a Second Language at Southern Illinois University. The normalized difference between pre- and post-test logits of difficulty (item response theory) and the pre-to-post difference index (criterion referencing) were found to be the most robust. Numerous problems, including spuriously high indices, were encountered with the Bayesian indices. It is concluded that decisions about discarding or revising test items should not be based solely on instructional sensitivity indices. Five data tables are included. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED301576

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

KYLE PERKINS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A Comparison of Instructional Sensitivity Indices

Kyle Perkins

Southern Illinois University

BEST COPY AVAILABLE

TM 012 435

A Comparison of Instructional Sensitivity Indices

Abstract

This paper presents a comparison of instructional sensitivity indices from four perspectives: criterion-referenced, classical theory, item response, and Bayesian. The normalized difference between pre and posttest logits of difficulty (item response theory) and the pre-to-post difference index (criterion-referenced) were found to be the most robust. Numerous problems, including spuriously high indices, were encountered with the Bayesian indices. The paper ends with a cautionary note about relying solely on instructional sensitivity indices in terms of discarding or revising test items.

This paper reports an attempt to estimate the instructional sensitivity of the items in a Michigan Test of English Language Proficiency grammar subtest and to compare the resulting indices of item effectiveness.

Haladyna and Roid (1981) defined instructional sensitivity as

> the tendency for (a test) item to vary in difficulty as a function of instruction. When instruction is reasonably effective, items given to uninstructed students should appear difficult and the very same items when administered to instructed students should appear easy. Items that do not detect differences from pretest to posttest occasions should be reviewed to ascertain if (a) instruction has been faulty or (b) the item is flawed in some way or (c) the item is inappropriate and does not belong in the domain. Instructional sensitivity is therefore a test concept which reflects the role played by instruction in affecting examinee performances on individual test items (p. 40).

For this research the following indices were used to estimate the instructional sensitivity of the Michigan grammar items: (1) the pre-to-post difference index (PPDI) introduced by Cox and Vargas (1966); (2) the percent of possible gain (PPG) introduced by Brennan and Stolurow (1971); (3) the combined-samples (instructed and uninstructed examinees) point-biserial correlation (COMPBI) introduced by Haladyna (1974); (4) four indices, 01, 11, 00, and 10, introduced by Popham (1971); (5) three test item discrimination indices, D1, D2, and D3, suggested by Helmstadter (1974); (6) item response indices: pretest logit of difficulty (PRELOGIT), posttest logit of difficulty (POSTLOGIT), pre-to-post difference in logit of difficulty (XLOGIT), and the normalized difference in item response difficulty estimates for the uninstructed and instructed samples (ZDIFF); and (7) three indices

based on the Bayes' theorem which were proposed by Helmstadter (1974) and Haladyna and Roid.

PPDI is the difference in item difficulty (determined as the proportion of correct responses) observed when an item is presented to uninstructed examinees, and then after they have been instructed. According to Roid and Haladyna (1982), "this index is the simple difference between the difficulty observed in the post-instruction group and the difficulty observed in the pre-instruction group. It therefore ranges from -1.00 to +1.00, with indexes of zero or lower quite rare. In most instructional settings, this index typically ranges from .10 to .60" (pp. 218-219).

PPG is a variation of PPDI, and it is computed by the formula

$$PPG = \frac{\text{difficulty (posttest)} - \text{difficulty (pretest)}}{1 - \text{difficulty (pretest)}}$$

PPG "is based on the notion that with any pretest item difficulty, there can only be a maximum gain, which is 1.00 minus the pretest difficulty" (Roid and Haladyna 1982: 219). This quote does not address the possibility that the posttest difficulty might be higher than the pretest difficulty, a phenomenon that can occur in English as a Second Language measurement.

The point-biserial correlation is a common procedure for computing item discrimability with point-biserial correlations of .25 and above being acceptable discrimination indices. Haladyna (1974) recommended combining uninstructed and instructed samples for the point-biserial correlation because more test-score variance is available with the combined samples, and COMPBI is less affected by range restriction. Haladyna noted that the COMPBI seems to function as a type of instructional sensitivity index.

The 01, 11, 00, and 10 indices were introduced by Popham (1971). When examinees take a pretest and a posttest, there are four logical states of of affairs: (1) 01, an incorrect response on the pretest and a correct

response on the posttest; (2) 11, correct responses on both the pre and posttest; (3) 00, incorrect responses on both tests; and (4) 10, a correct response on the pretest and an incorrect response on the posttest. The percentage in the 01 cell for examinees who took both tests would provide an index of instructional sensitivity.

Roid and Haladyna (1982) provided a fourfold table to represent the four possible patterns of item responses from the pre and posttest:

	Posttest	
Pretest	Correct	Incorrect
Incorrect	01	00
Correct	11	10

(p. 220)

D1, D2, and D3 are three separate variants of a classical item discrimination index which is computed using the sample separation techniques. These three variants involve defining "high knowledge" and "low knowledge" groups in different ways. The following descriptions came from Helmstadter (1974).

D1 assumed that persons in the top one-third of the class on the posttest were "high knowledge" and persons in the bottom one-third of the class on the posttest were "low knowledge."

D2 involved combining scores from the pre and posttest as if they constituted one large class and then assuming that persons in the top one-third of this doubled class were "high knowledge" and that persons in the bottom one-third of this class were "low knowledge."

D3 assumed that the pretest scores represented a "low knowledge" group and that posttest scores represented a "high knowledge" group.

The application of the Rasch one-parameter item response model produced item characteristics which were considered in this study. The

pre-to-posttest difference in logit of difficulty (XLOGIT) and its normalized equivalent, z-difference (ZDIFF), were considered for this research as measures of instructional sensitivity. Haladyna and Roid (1981) obtained a high degree of correspondence between ZDIFF and other instructional sensitivity indices.

Helmstadter (1974) and Haladyna and Roid (1981) proposed three instructional sensivity indices which were based on the Bayes' theorem (Bayes, 1763). Bayes, an English clergyman who died in 1760, produced a formal derivation of the equality among certain probabilities, and the theorem "is based on the fact that the joint probability of two events A and B can be written as the probability of one of the events and conditioned probability of the second events, given the first event" (Iverson, 1984, p. 12). Bayes' theorem relates conditional and marginal probabilities as follows:

$$P(B|A) = P(A|B)\ P(B)/P(A).$$

The three indices, based on Bayes' theorem and proposed by Haladyna and Roid, require collateral information; the most appropriate type of information is preinstruction test results. The three Bayesian indices are: (a) B1, the probability that an examinee knows the content material given that the correct response was selected; (b) B2, the probability that an examinee does not know the content material given that the incorrect was selected; and (c) B3, the probability that a correct decision will be made about the examinee's knowledge of the content given the results of performance on that item, that is, mastery or nonmastery. An implication is that items with high B3 indices augment the test administrator's probability of correctly identifying mastery and nonmastery students.

Haladyna and Roid (1981) provided the Bayesian indices' computing formulas which entail pretest and posttest item difficulties:

$$B1 = \frac{POSTDIFF*COMDIFF}{(POSTDIFF*COMDIFF) + (PREDIFF*[1-COMDIFF])}$$

$$B2 = \frac{(1-PREDIFF) * (1-COMDIFF)}{([1-PREDIFF) * (1-COMDIFF)] + ([1-POSTDIFF) * (COMDIFF)]}$$

$$B3 = (1-COMDIFF) * (1-PREDIFF) + (COMDIFF*POSTDIFF)$$

Where

B1, B2, and B3 are the three Bayesian indexes, and PREDIFF is the preinstruction sample difficulty, POSTDIFF is the postinstruction sample difficulty, and COMDIFF is the mean of PREDIFF + POSTDIFF, or the combined-samples' difficulty (p. 43).

Thomas Haladyna (personal communication) has supplied the following conceptual bases for each index:

p denotes a correct response.

$\bar{p}$ denotes an incorrect response.

k denotes the condition where the examinee has knowledge (mastery).

$\bar{k}$ denotes the condition where the examinee does not have knowledge (nonmastery).

P(k) is the probability of a correct response given that the examinee has knowledge (COMDIFF).

$P(p|\bar{k})$ is the probability of a correct response given that the examinee is a nonmaster (PREDIFF).

P(p|k) is the probability of a correct response given that the examinee is a master (POSTDIFF).

$P(\bar{k})$ is the probability that the examinee is a nonmaster (1.00-COMDIFF).

$P(\bar{p}|\bar{k})$ is the probability of an incorrect response given that the examinee is a nonmaster (1.00-PREDIFF).

$P(\bar{p}|k)$ is the probability of an incorrect response given that the examinee is a master (1.00-POSTDIFF).

P(k|p) is the probability that the examinee is a master given a correct response for the item, which is B1 and is equal to

$$\frac{P(p|k)\ P(k)}{P(p|k)\ P(k) + P(p|\bar{k}) * P(\bar{k})}$$

or

$$\frac{POSTDIFF * COMDIFF}{(POSTDIFF * COMDIFF) + (PREDIFF * (1\text{-}COMDIFF))}$$

$P(\bar{k}|\bar{p})$ is the probability that the examinee is a nonmaster given an incorrect response for the item, which is B2 and is equal to

$$\frac{P(\bar{p}|\bar{k})\ P(\bar{k})}{P(\bar{p}|\bar{k})\ P(\bar{k}) + P(\bar{p}|k)\ P(k)}$$

or

$$\frac{(1\text{-}PREDIFF) * (1\text{-}COMDIFF)}{((1\text{-}PREDIFF) * (1\text{-}COMDIFF)) + ((1\text{-}POSTDIFF) * (COMDIFF))},$$

$P(k\&p) + P(k\&p)$ is the probability of a correct decision (mastery-nonmastery which is B3 and is derived thusly:

$P(\bar{k}$ and $\bar{p})$ or $P(\bar{k})\ P(\bar{p}|\bar{k}) + p(k$ and $p)$ or $P(k) * P(p|k)$

or

$(1\text{-}COMDIFF) * (1\text{-}PREDIFF) + (COMDIFF * POSTDIFF)$.

In a further personal communication, Haladyna noted that any B1, B2, and B3 index greater than .70 would reflect reasonable instructional sensitivity. This recommendation comes from a consideration of reasonable pre-to-post difference indices and converting them to the equivalent Bayesian indices.

I Method

a Subjects: This research was conducted at the Center for English as a Second Language at Southern Illinois University. The subjects who participated in this study were enrolled full-time in proficiency level 3 at the Center. Placement of students into the four full-time proficiency levels (1, 2, 3, and 4) is determined by the results of an Institutional TOEFL test which is administered to new students

at the beginning of each term. Students who score between 430 and 469 are placed into level 3.

The distribution of subjects by native language was the following: Arabic, 5; Farsi, 2; French, 1; Fula, 1; Greek, 1; Japanese, 2; Korean, 2; Portuguese, 1; Spanish, 2; Wolof, 1.

b Material and c Procedure: At the beginning of an eight-week term of instruction, the examinees were given a Michigan Test of English Language Proficiency grammar subtest. The examination was administered under test conditions. Working time was thirty minutes. The examinees were told that they would be tested again with the same test at the end of the term. During the last week of the eight-week term the same test was administered again under the same conditions with thirty minutes working time. The grammar class met five days per week, fifty minutes per day for eight weeks. The textbook used in level 3 was Harris's (1981) Practical English 3.

II. Results and discussion

(Put Table 1 about here)

Table 1 presents the means and standard deviations for the two administrations of the Michigan grammar test. The difference between the two means was significant: $t=4.39$, $p<.01$, df 17, one-tailed test. The correlation obtained between the pre and posttest scores was also significant: $r=.629$, $p<.01$, df 16, one-tailed test.

The data were submitted to the various analyses presented above. Two notes of explanation are necessary concerning the Rasch one parameter item response theory model. A Rasch analysis was incorporated into this study because the Rasch model yields estimates of difficulty for test items and estimates of ability for persons. Item difficulty and person ability are expressed in logits which employ a true interval scale with mean difficulty set at zero. Logits are adjusted for

sample ability spread and test width in order to eliminate boundary effects and skewing. Negative values represent low-difficulty items and low-ability persons; positive values represent high-difficulty items and high-ability persons.

There were no pretest logits of difficulty for two items because, on the pretest, there were no correct responses for these items. The Rasch model must have variability for both item and person scores; logits cannot be estimated for items with difficulties of .00 or 1.00, or for person (test) scores of .00 or 1.00 (cf. Wright and Stone, 1979).

This subject pool mastered six items, or 15 percent of the test. Three items had extremely low pretest difficulties and spuriously high B1 indices. One can infer that extremely low pretest item difficulties produce excessively high B1 indices. Five items exhibited low (negative logits) pretest difficulties but high (positive logits) posttest difficulties, that is, this subject pool showed poorer posttest performance on five items than pretest performance.

Eight of the structures tested were not taught during the eight-week term. Table 2 presents the item difficulties, logits, and B1 indices for each of the categories of items discussed above.

(Put Table 2 about here)

(Put Table 3 about here)

Table 3 presents the ranges and medians for the various indices.

(Put Table 4 about here)

After the indices were calculated for the forty grammar items (thirty eight items for the Rasch data), they were submitted to Pearson product moment correlation. The criterion level of significance was set at the .01 level in advance. A 17 x 17 correlation matrix

is too large to reproduce in this format; therefore, Table 4 lists only the significant correlations for each index.

The significant correlations for which each index entered as a variable were averaged to estimate the robustness of each index. The Fisher Z transformation was utilized to transform the correlation coefficients from a distorted ordinal scale to a normal interval scale before the averages were calculated. After the means were calculated, they were converted back to correlation coefficients. The mean significant correlation coefficients and the number of significant correlations for each index are reported in Table 5.

(Put Table 5 about here)

As Tables 4 and 5 indicate, the more robust indices of instructional sensitivity were ZDIFF and PPDI; robust is used in this context to refer to the magnitude of the average correlation and the number of significant correlations entered.

Haladyna and Roid (1981) reported high correlations between ZDIFF and other instructional sensitivity indices. In addition, ZDIFF's robustness is probably attributable to its being sample independent and to its providing normalized scale values.

PPDI has several positive attributes. First, it was consistently related to many of the other indices; similar findings have been reported for first language, non-language test data by Haladyna and Roid (1981) and Herbig (1976). Second, PPDI is the easiest index to calculate and is, therefore, highly recommended to classroom teachers. It has one obvious drawback: it requires pre and posttest data.

The Bayesian indices have many vexing problems. In a previous section it was noted that low pretest difficulties can generate spuriously high B1 indices. Haladyna and Roid (1981) have identified further problems:

> The family of Bayesian indexes has enough problems to be eliminated from all but further research use. B1 is too easily influenced by high pretest and posttest difficulties; B2 is too easily influenced by low pretest and posttest difficulties; and while B3 is the most dependable of the three, it too is influenced by ceiling and floor effects. It is not recommended that any of the Bayesian indices be used for these reasons, and, further, since Bayesian indexes require complex computation (p. 50).

In conclusion, the purpose of this paper was to compare instructional sensitivity indices from criterion-referenced testing, classical test theory, item response theory, and Bayesian perspectives. ZDIFF and PPDI were found to be robust indices; however, on the subject of instructional sensitivity indices in general, the following quote from Haladyna and Roid (1981) is a fitting coda:

> Neither a high ISI (instructional sensitivity index) nor a high difficulty index can conclusively determine whether to keep, discard or revise an item. These data must be used in conjunction with evidence from a logical review. Without a logical review of items and their relationship to instruction, the empirical item review becomes a meaningless exercise. The existence of a low ISI or difficulty should not lead directly to the revision or discarding of an item, because it may be that instruction was faulty, not the item's ability to detect instructional effects (p. 51).

References

Bayes, T. 1763. Essay towards solving a problem in the doctrine of chances. Philosophical Transactions 53, 370-418. (Reprinted in Biometrika, 1958, 45, 293-315).

Brennan, R.L. and Stolurow, L.M. 1971. An empirical decision process for formative evaluation. Research Memorandum No. 4. Cambridge, MA: Harvard University CAI Laboratory.

Cox, R.C. and Vargas, J. 1966. A comparison of item selection techniques for norm-referenced and criterion-referenced tests. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, April.

Haladyna, T. 1974. Effects of different samples on item and test characteristics of criterion-referenced tests. Journal of Educational Measurement 11, 93-100.

Haladyna, T. and Roid, G. 1981. The role of instructional sensitivity in the empirical review of criterion-referenced test items. Journal of Educational Measurement 18, 39-52.

Harris, T. 1981. Practical English 3. New York: Harcourt Brace Jovanovich, Inc.

Helmstadter, G.C. 1974. A comparison of Bayesian and traditional indexes of test item effectiveness. Paper presented at the annual meeting of the National Council of Measurement in Education, Chicago, April.

Herbig, M. 1976. Item analysis by use in the pretests and posttests: A comparison of different coefficients. Programmed Learning and Educational Technology 13, 49-54.

Iversen, G.R. 1984. Bayesian statistical inference. (Sage University paper No. 43. Series: Quantitative Applications in the Social Sciences). Beverly Hills, CA: Sage Publications, Inc.

Popham, W.J. 1971. Indices of adequacy for criterion-referenced tests. In Popham, W. J., editor, Criterion-referenced measurement. Englewood Cliffs, N.J.: Educational Technology Publications.

Roid, G. and Haladyna, T. 1982. A technology for test-item writing New York: Academic Press.

Wright, B.D. and Stone, M.H. 1979. Best test design: Rasch measurement. Chicago: MESA Press.

Table 1

Means and Standard Deviations

Test	No. of subjects	No. of items	Mean	Standard deviation
Pre	18	40	17.00	2.27
Post	18	40	21.28	4.46

$t = 4.39$, $p < .01$, df 17, one-tailed test

Table 2

Categories of items

Mastery items with B1's greater than .70

Item No.	Pretest difficulty	Posttest difficulty	Pretest logit	Posttest logit	B1
1	.33	.72	+0.43	-0.85	.71
3	.56	.83	-0.59	-1.53	.77
11	.33	.72	+0.43	-0.85	.71
19	.78	.78	-1.71	-1.20	.78
33	.72	.83	-1.35	-1.53	.80
39	.22	.72	+1.04	-0.85	.74
		Spuriously high B1 indices			
6	.00	.56	N/A	-0.12	1.00
37	.06	.50	+2.64	+0.14	.76
38	.00	.11	N/A	+2.33	1.00
		Low pretest but high posttest difficulty items			
16	.44	.39	-0.08	+0.61	.39
17	.67	.50	-1.10	+0.14	.52
28	.56	.39	-0.59	+0.61	.39
29	.50	.39	-0.33	+0.61	.39
32	.56	.33	-0.59	+0.88	.33

Structures not taught in the syllabus

Item No.	Pretest difficulty	Posttest difficulty	Pretest logit	Posttest logit	B1
5	.61	.56	-0.82	-0.12	.57
14	.44	.56	-0.08	-0.12	.56
18	.56	.56	-0.59	-0.12	.56
20	.33	.33	+0.43	+0.88	.33
24	.67	.56	-1.10	-0.12	.58
34	.06	.33	+2.64	+0.88	.58
35	.50	.56	-0.33	-0.12	.56
40	.06	.17	+2.64	+1.81	.28

Table 3

Ranges and medians

Index	Range	Median
PPDI	-.23 to .56	.06
PPG	-.52 to .64	.12
ZDIFF	25.27 to 65.54	52.91
COMPBI	-.47 to .55	.18
01	.06 to .56	.17
00	.00 to .89	.30
11	.00 to .72	.33
10	.00 to .33	.11
D1	.20 to 1.00	.62
D2	.27 to 1.00	.63
D3	.38 to 1.00	.54
PRELOGIT	-1.71 to +2.64	-.08
POSTLOGIT	-1.53 to +2.33	-.12
XLOGIT	-2.50 to +1.47	+.23
B1	.20 to 1.00	.58
B2	.20 to .95	.57
B3	.39 to .95	.58

Table 4

Significant correlations

PPDI and		ZDIFF and		01 and	
ZDIFF	-.968	D3	-.969	PPDI	.895
01	.895	PPDI	-.968	ZDIFF	-.825
XLOGIT	-.836	B3	-.875	XLOGIT	-.753
D3	.818	01	-.825	PPG	.669
PPG	.782	XLOGIT	.809	D3	.635
B3	.757	PRELOGIT	-.762	B3	.544
B2	.667	B2	-.752	B2	.499
10	-.664	PPG	-.714	11	-.493
PRELOGIT	.635	10	.707	PRELOGIT	.474
COMBI	.536	11	.582	B1	.468
11	-.525	COMPBI	-.542	COMPBI	.44
B1	.523	B1	-.402		

PPG and		COMPBI and		00 and	
PPDI	.782	D2	.718	POSTLOGIT	.896
ZDIFF	-.714	D1	.555	PRELOGIT	.725
01	.669	ZDIFF	-.542	B2	.715
XLOGIT	-.643	PPDI	.536	11	-.669
B3	.593	PPG	.518	10	-.510
D3	.566	B3	.489	D3	.457
10	-.539	D3	.462	B3	.414
B1	.526	01	.444		
COMPBI	.518	10	-.416		
POSTLOGIT	-.403				

11 and	
B2	-.922
PRELOGIT	-.905
D3	-.713
00	-.669
POSTLOGIT	-.635
ZDIFF	.582
B3	-.582
PPDI	-.525
01	-.493

D1 and	
D2	.861
COMPBI	.555

D2 and	
D1	.861
COMPBI	.718

D3 and	
ZDIFF	-.969
B3	.919
PRELOGIT	.847
PPDI	.818
B2	.817
10	-.739
11	-.713
XLOGIT	-.712
01	.635
PPG	.566
B1	.534
COMPBI	.462
00	.437

PRELOGIT and	
B2	.959
11	-.905
D3	.847
ZDIFF	-.762
B3	.726
00	.725
PPDI	.635
10	-.597
POSTLOGIT	.539
01	.474
XLOGIT	-.434

XLOGIT and

PPDI	-.836
ZDIFF	.809
01	-.753
D3	-.712
PPG	-.643
B3	-.623
10	.544
B1	-.531
B2	-.463
PRELOGIT	-.434

B1 and

D3	.534
XLOGIT	-.531
PPG	.526
PPDI	.523
POSTLOGIT	-.501
B3	.488
01	.468
ZDIFF	-.402

B2 and

PRELOGIT	.959
11	-.922
D3	.817
ZDIFF	-.752
B3	.743
00	.715
PPDI	.667
10	-.639
PPDI	.667
10	-.639
POSTLOGIT	.539
01	.499
XLOGIT	-.463

B3 and

00	.414
D3	.919
ZDIFF	-.875
PPDI	.757
10	-.754
B2	.743
PRELOGIT	.726
XLOGIT	-.623
PPG	-.593
11	-.582
01	.544
COMPBI	.489
B1	.414

10 and

B3	-.754
D3	-.739
ZDIFF	.707
PPDI	-.664
B2	-.639
PRELOGIT	-.597
XLOGIT	.544
PPG	-.539
OO	-.510
COMPBI	-.416

Table 5

Mean Significant Correlations

Index	Mean significant correlation	Number of significant correlations
ZDIFF	.800	12
D2	.800	2
PPDI	.760	12
D3	.755	13
B2	.745	11
D1	.743	2
PRELOGIT	.740	11
11	.710	9
B3	.685	13
00	.660	7
XLOGIT	.652	10
01	.640	11
10	.620	10
POSTLOGIT	.620	6
PPG	.604	10
COMPBI	.525	9
B1	.495	8